

**Hyper text-to-speech conversion method****Publication number:** TW434492 (B)**Publication date:** 2001-05-16**Inventor(s):** JUNG JIN-JIUN [TW]; HUANG SHAU-HUA [TW]; JUNG CHUNG-BIN [TW]**Applicant(s):** IND TECH RES INST [TW]**Classification:****- international:** G06F17/00; G06F17/00; (IPC1-7): G06F17/00; G10L7/02**- European:****Application number:** TW19980110229 19980625**Priority number(s):** TW19980110229 19980625**Abstract of TW 434492 (B)**

The present invention relates to a system that can convert hyper text into speech signals. The system consists of a hyper text markup language (HTML) analyzer, an articulation control command analyzer, a tag converter, a text converter, and a traditional text-to-speech converter. The HTML analyzer reads and analyzes the input hyper text and divides it into a text content, a HTML tag for marking up the text structure, and an articulation control command for controlling the way of articulation. The articulation control command analyzer analyzes the articulation control command and, in accordance with the command type, stores the command content into a tag corresponding table, a sound effect table, a parameter table, a heteronym table, or a proper noun table.; In accordance with the heteronym table and the proper noun table, the text converter carries out text replacement or conversion to correct the text articulation. The tag converter not only analyzes the HTML tag and, in accordance with the tag corresponding table and the parameter table, controls the traditional text-to-speech converter's way of articulation but also, in accordance with the tag corresponding table and the sound effect table, synthesizes the sound effects to allow listeners to easily distinguish the text structure. Based on the way of articulation set by the tag converter, the traditional text-to-speech converter converts the text content converted by the text converter into speech signals.

---

Data supplied from the *esp@cenet* database — Worldwide

# 公告本

申請日期	87.6.25
案 號	87110229
類 別	G06F 17/00, G10L 7/02

A4  
C4

434492

(以上各欄由本局填註)

發 明 專 利 說 明 書		
一、發明 新 型	中 文	一種超媒體文件轉換為語音之方法
	英 文	
二、發明 創 作 人	姓 名	鍾錦鈞 黃紹華 鍾崇斌
	國 籍	中華民國 中華民國 中華民國
三、申請人	住、居所	新竹縣新埔鎮鹿鳴里6鄰53號 新竹市東區光明里14鄰大學路88號10樓之2 新竹市東區軍功里34鄰建功一路70巷7號5樓
	姓 名 (名稱)	財團法人工業技術研究院
	國 籍	中華民國
	住、居所 (事務所)	新竹縣竹東鎮中興路4段195號
	代 表 人 姓 名	孫 震

裝

訂

線

經濟部中央標準局員工消費合作社印製

## 四、中文發明摘要(發明之名稱：一種超媒體文件轉換為語音之方法)

一種將超媒體文件轉換為語音訊號的系統，此系統包含一超媒體標示語言剖析器、一發音控制指令剖析器、一標籤轉換器、一文字轉換器及一傳統的文字語音轉換器。超媒體標示語言剖析器讀取並剖析輸入的超媒體文件，將其分離成文字內容、標示文件結構的超媒體標籤及控制發音方式的發音控制指令。發音控制指令剖析器分析發音控制指令，並依據指令的類別將其內容存入標籤對應表、音效表、參數表、破音字表或專有名詞表中。文字轉換器依據破音字表及專有名詞表，執行文字的替代轉換以更正文字的發音。標籤轉換器分析超媒體標籤，並依據標籤對應表及參數表，控制傳統文字語音轉換器的發音方式。同時標籤轉換器也依據標籤對應表及音效表執行音效的合成，讓聽眾易於分辨文件的結構。傳統文字語音轉換器則依據標籤轉換器所設定的發音方式，將經由文字轉換器轉換完成的文字內容，轉換為語音訊號。

## 英文發明摘要(發明之名稱：

## 五、發明說明( / )

### 背景說明

文字語音轉換器(text-to-speech converter)是一種將文字轉換為語音的裝置。對有視覺障礙的人而言，這種裝置可幫助他們聽取外界的資訊。在特定的環境下，這種裝置也是一般人獲取資訊的一種重要配備，例如開車的時候，或是使用電話的時候。而這些資訊的來源，可能是電子文件，或是透過光學掃描器及文字辨識裝置所轉換而來的文字資訊。

在日常生活中，電子式資訊之來源日益眾多，且成級數之成長，例如電子郵件、行事曆、電子新聞、股票資訊及備受矚目的全球資訊網。要將這些電子資訊轉換為數位語音，若採用人工錄音再加以數位化處理，不僅需要耗費龐大的人力及巨大的儲存空間，而且人工錄音的方式，無法適用於電腦系統依使用者需求，自動彙整產生的電子資訊。

如何將各種原本適用於視覺顯示的電子資訊轉換為語音型式，對文字語音轉換器的設計人員而言，確實是一種挑戰。其中最主要的原因在於電子資訊的呈現，不僅需要呈現其文字內容，同時也需要考慮這些文字內容的呈現方式，例如在視覺顯示中的大小寫、粗體、斜體、段落型式及列舉型式等呈現方式。在進行文字語音轉換時，這些原本用於控制視覺顯示的格式及字體控制碼，不能直接轉換為語音。在文字內容中的標點符號，也是不能直接轉換為語音。另外，一字串的發音，在不同的前後文，也會有不同的發音方法，例如中文的破音字發

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

東

## 五、發明說明(一)

音，就是一典型的例子。為了解決這些問題，先前的技藝發明提出了各種的解決方法。

美國專利第 5,555,343 號揭示一種解決這類問題的文字語音轉換技術。其中包含格式化及字體控制碼的處理方式，以及標點符號及特定的文數字格式的處理方法。此方法採用第一事先建好的表格，將格式化及字體控制碼對應成語音控制碼，用來控制發音之速度或音量的大小。此方法採用第二個事先建好的表格，將特定的文數字格式對應成口語化的文字字串。這些特定的文數字格式包括用來表示時間而以冒號分隔的數字字串、用來表示日期而以斜線分隔的數字字串及用來表示檔案目錄而以斜線分隔的文字字串等。此方法採用第三個事先建好的表格，將標點符號或數學運算符號對應成口語化的文字字串或語音控制碼。此方法使用一事先建好的表格，以決定一輸入的字元為可發音或是不可發音。遇到不可發音的字元，才依前述的第一、第二個及第三個事先建好的表格，決定適當的發音方式。

美國專利第 5,634,084 號揭示另一種解決這類問題的文字語音轉換技術。此方法先將輸入之文字，依上下文的關係，加以分類成數字、度量單位、地理名詞及時間日期等類別。再將此分類後之文字，依一或多個不同類別的縮寫字表加以展開，對應成口語化之字詞。例如地名的縮寫“SF, CA”，此方法可將其轉換為“San Francisco California”；亦可將“MPEG”轉換為口語化的“m peg”。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明(3)

由於網際網路(Internet)及全球資訊網(World Wide Web)的普及，全球資訊網已成為現今電子資訊的主要來源之一。全球資訊網上的電子資訊，大部份是採用超媒體標示語言(Hyper Text Markup Language, HTML)的格式，我們稱之為超媒體文件。超媒體文件與其他電子文件不一樣的地方，是在其原始文件中，除了文件內容外，還含有超媒體標籤(HTML tag)。超媒體標籤是超媒體標示語言所定義的文字標籤，用於標示文件之內容與結構，或是文件之顯示控制。例如下面這個例子表示一段超媒體文件的原始文件：

```
<!BODY BGCOLOR=#DBFFFF>
<body bgcolor=white>
<CENTER>
<map name="Main">
  <area shape="rect" coords="157,12,257,112"
    href="Main.html">
  <area shape="rect" coords="293,141,393,241"
    href="VRML.html">
  <area shape="rect" coords="18,141,118,241"
    href="VRML.html">
  <area shape="rect" coords="157,226,257,366"
    href="Main.html">
</map>

```

## 五、發明說明(4)

border=0></img>

<br><br><br><br>

<b>

<font size=3 color=black>

Welcome to the VR workshop of our company

</font>

<a href="http://www.ccl.itri.org.tw" ><font size=3

color=blue>ITRI</font></a>

<font size=3 color=black>/</font>

<a href= "http://www.ccl.itri.org.tw"><font size=3

color=blue>CCL</font></a>

<font size=3 color=black>. We have been<br>

developing some advanced technologies as follows,<br>

</b>

<ul>

<a href="Main.html">

<li><font size=3 color=blue> PanoVR</font>

</a>

<font size=3>(A panoramic image-based VR)</font><br>

<a href="VRML.html">

<li><font size=3 color=blue>CyberVR</font>

</a>

<font size=3>(A VRML 1.0 browser)</font><br>

</ul>

<br><br>

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

象

## 五、發明說明(5)

```

<a href="Winner.html"></img></a><br>
</a>
<br><br>
<font size=3 color=black>
<br>You are the th
visitor<br>
</font>
<HR size=2 WIDTH=480 ALIGN=CENTER>
(C) Copyright 1996 Computer and Communication
Laboratory,<br>
Industrial Technology Research Institute, Taiwan, R.O.C.
</BODY>

```

由這個例子可看到，在超媒體文件的原始文件中，全都是可顯示的字元，沒有特殊無法顯示的控制碼。超媒體標籤是以“<”及“>”標示，並分為起始標籤及結束標籤。起始標籤以“<”起頭，而結束標籤則以“</”起頭。因此，“<font size=3 color=black>”是“font”超媒體標籤的起始標籤，而“</font>”則為其結束標籤。

超媒體標示語言將由超媒體標籤的起始標籤及結束標籤所標示的文字內容賦予特殊的意義，以表達文件的結構，例如標題、段落、列舉及表格等。而這些結構元

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線



## 五、發明說明(6)

件的顯示方式，則是由全球資訊網瀏覽器所控制。因此，同一超媒體文件，在不同的全球資訊網瀏覽器或顯示系統上，會有不同的顯示方式。另外超媒體標籤也允許巢狀的排列方式，例如在上例中，“<b>”及“<font size=3 color=black>”以巢狀的方式加諸於“Welcome to the VR workgroup of our company”。

雖然上述範例是以英文撰寫的超媒體文件，但超媒體文件的文字內容允許用其他的語言，如中、日及韓文等。對一般全球資訊網中文網頁的讀者而言，對一些英文的專有名詞，如“Web”、“World Wide Web”及“HTML”等名詞，並不陌生。因此在一般的中文網頁中，常會混雜英文的專有名詞。另一方面，在以超媒體標示語言所撰寫的技術文件、開會通知及備忘錄等文件中，也是常使用英文的專有名詞或縮寫。因此在一般的全球資訊網網頁中，常混雜有多種的語言。除此之外，我們還需考慮同一字串具有多種發音的問題，例如中文的破音字，同一字在不同的詞或前後文中，會有不同的發音方式。

上述傳統的文字語音轉換器並不適用於將超媒體文件轉換為語音訊號。在轉換的過程中，首先必須解決的問題，就是如何剖析超媒體文件以辨識超媒體標籤。由於超媒體標籤均由可顯示的字元所組成，而非特殊的控制碼，而且超媒體標籤允許巢狀式的排列方式，因此先前有關文字語音轉換器的技藝發明並不適用於剖析超媒體文件。再者，上述傳統的文字語音轉換器並無法解決破音字問題。雖然有些中文文字語音轉換器可解決部份

(請先閱讀背面之注意事項再填寫本頁)

發

訂

線

## 五、發明說明(7)

的破音字問題，但這些文字語音轉換器並未考慮多種語言混雜的問題。本發明的重點即在克服這些問題。

### 發明概要

依據本發明所提出的方法，可以將超媒體文件轉換為語音訊號，或是達到其他的目的。在此所提出的實施例，是一將超媒體文件轉換為語音訊號的電腦系統。這個電腦系統包含一超媒體標示語言剖析器、一發音控制指令剖析器、一標籤轉換器、一文字轉換器及一傳統的文字語音轉換器。超媒體標示語言剖析器讀取並剖析輸入的超媒體文件，將其分離成文字內容、標示文件結構的超媒體標籤及控制發音方式的發音控制指令。發音控制指令剖析器分析發音控制指令，並依據指令的類別將其內容存入標籤對應表、音效表、參數表、破音字表或專有名詞表中。文字轉換器依據破音字表，將超媒體文件的文字內容中，凡是必須修改發音的字串，以替代字串修正文字的發音。同時文字轉換器也依據專有名詞表，將超媒體文件的文字內容中，凡是必須翻譯的字串，以譯文字串取代原字串的發音。標籤轉換器分析超媒體標籤，並依據標籤對應表及參數表，控制傳統文字語音轉換器的發音參數，以改變該超媒體標籤所標示的文字內容的音量、速度及韻律等發音參數。同時標籤轉換器也依據標籤對應表及音效表執行音效的合成，讓聽眾易於分辨文件的結構。傳統文字語音轉換器則負責將經由文字轉換器轉換完成的文字內容，依據標籤轉換器所設定的發音方式，轉換為語音訊號。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明(8)

這個依據本發明所建構的系統，可以將超媒體文件轉換為語音訊號，同時也解決破音字及語言混雜等問題。這個系統不但設計精巧，同時也極具應用與擴充能力。此裝置與方法不但可提供個人化的文字語音轉換，而且也讓超媒體文件的提供者，更易於設計超媒體文件的語音表達方式。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

束

## 五、發明說明(9)

### 圖式簡單說明

圖 1 說明本發明的一具體實施例。

圖 2 說明本發明各個元件間的資料流程。

圖 3 說明一序列隱藏於超媒體文件中的發音控制指令。

圖 4A、4B 及 4C 分別說明參數表、音效表及標籤對應表。

圖 5A 及 5B 分別說明破音字表及專有名詞表。

圖 6 說明文件讀取控制器的執行步驟。

圖 7 說明發音控制指令剖析器的執行步驟。

圖 8 說明文字轉換器的執行步驟。

圖 9 說明標籤轉換器的執行步驟。

### 發明的詳細說明

圖 1 說明一依據本發明所建構的超媒體語音轉換系統 10。此系統是一電腦系統，內含一中央處理器 11、主記憶體 12、網路裝置 13、電話介面裝置 14、鍵盤和滑鼠 15、音效裝置 16、顯示器 17 以及儲存裝置 18。此系統使用一條排線 19 將這些裝置 11-18 連接在一起。經由此排線 19，這些裝置 11-18 間可傳輸指令或資料。

儲存裝置 18 可以是磁碟機，用於儲存資料及程序(process)。主記憶體 12 也是儲存資料及程序，但通常是用來儲存目前中央處理器 11 正在執行的指令及資料。中央處理器 11 是用來執行程序中的指令並處理資料。網路裝置 13 是用來與電腦網路連接，例如乙太網

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

象

## 五、發明說明(10)

路連接器或是其他型態之網路卡。電話介面裝置 14 是用來與電話網路連接。鍵盤和滑鼠 15 是用於接收使用者輸入的命令或資料。顯示器 17 是以文字型式或圖形型式顯示電子資訊。音效裝置 16 接收數位語音訊號，並透過喇叭或耳機發出聲音或音效。

如圖 1 所示，儲存裝置 18 內儲存了作業系統、應用程式、超媒體文件檔案 23、發音控制指令檔 21 及文件讀取模組 29。作業系統及應用程式是一般熟知的技術，在此不再贅述。文件讀取模組 29 包含文件讀取控制器 28、文字語音轉換器 27、超媒體標示語言剖析器 24、發音控制指令剖析器 22、標籤轉換器 25、標籤對應表 41、參數表 42、音效表 43、文字轉換器 26、破音字表 31 及專有名詞表 32。

雖然上述之程序 22，24-28 是以分時共享的方式，由中央處理器 11 負責執行，但這只是為了方便陳述本發明之方法。上述之程序 22，24-28 也可使用熟知的硬體技術，以硬體的實施方式來達到相同的功能，在此不再贅述這種實施方式。另外文字語音轉換器 27 及超媒體標示語言剖析器 24 也是熟知的技藝，在此不再詳述。

文件讀取控制器 28 負責控制整個轉換過程中，各個程序 22，24-27 間的資料流程。圖 2 顯示文件讀取控制器 28 內之資料流程。超媒體文件檔案 23 可能源自於網路裝置 13，或是從儲存裝置 18 讀取。發音控制指令檔 21 也是源自於網路裝置 13，或是從儲存裝置 18 讀取。

(請先閱讀背面之注意事項再填寫本頁)

派

訂

發

## 五、發明說明 ( || )

超媒體標示語言剖析器 24 分析超媒體文件檔案 23 的內容，將其分離成文字內容、標示文件結構的超媒體標籤及控制發音方式的發音控制指令。超媒體標示語言剖析器 24 將分離出的超媒體標籤送至標籤轉換器 25。超媒體標示語言剖析器 24 將分離出的文字內容送至文字轉換器 26。超媒體標示語言剖析器 24 將分離出的發音控制指令送至發音控制指令剖析器 22。

發音控制指令剖析器 22 負責分析發音控制指令。發音控制指令可能儲存於獨立的發音控制指令檔 21，或是隱含在超媒體文件檔案 23 中。發音控制指令分成下列四種型態：

- (1) 聲音控制指令，其格式為：PARAM 超媒體標籤 屬性 發音參數；
- (2) 音效控制指令，其格式為：AUDIO 超媒體標籤 屬性 音效資料；
- (3) 破音字控制指令，其格式為：ALT 破音字串 替代字串 前後文字串；
- (4) 專有名詞控制指令，其格式為：TERM 專有名詞字串 譯文字串。

圖 3 顯示一序列的發音控制指令 110, 120, 130, 140, 150, 160, 170 及 180。這些發音控制指令是使用超媒體標示語言中的註解標籤(comment tag)加以標示，使其能隱藏於超媒體文件中。指令 110 是一聲音控制指令，因其使用“PARAM”識別碼 111。這個聲音控制指令 110 定義所有以超媒體標籤 113 “LI”標示的文字內容，均

(請先閱讀背面之注意事項再填寫本頁)

衣

訂

線

## 五、發明說明 (12)

依據其發音參數 115 所定義的參數作聲音控制。這個聲音控制指令 110 中，發音參數 115 所定義的參數為：速度(speed)是 1.0，音量(volume)是 0.8，韻律(pitch)是 1.2。在聲音控制指令中，可以在超媒體標籤 113 及發音參數 115 間，選擇性地加入屬性欄位。屬性欄位是用於限制該聲音控制指令 110 所適用之範圍。例如屬性欄位可以是超媒體標籤的一些屬性，使用屬性欄位則可限制該聲音控制指令 110 只適用於據有特定屬性的超媒體標籤 113。

指令 120 是一音效控制指令，因其使用 "AUDIO" 識別碼 121。這個音效控制指令 120 定義碰到超媒體標籤 123 "LI" 時，必須插入音效資料 125 "beep.au"。在此例中，音效資料 125 是一取名為 beep.au 的音效資料檔。在音效控制指令 120 中，也可以選擇性地加入屬性欄位，用來限制此音效控制指令 120 所適用之範圍。

發音控制指令剖析器 22 在分析聲音控制指令及音效控制指令時，會依據指令的內容修改如圖 4A 所示的參數表 42，或是如圖 4B 所示的音效表 43。然後發音控制指令剖析器 22 再修改如圖 4C 所示的標籤對應表 41。

如圖 4A 所示，發音控制指令剖析器 22 在分析聲音控制指令 110 時，會在參數表 42 中，加入或修改項目 42-1。首先發音控制指令剖析器 22 必須在參數表 42 取得可以使用的項目 42-1。項目 42-1 的取得方式可以是在參數表 42 中插入新的項目，或是重覆使用參數表 42 中不再使用的項目。在取得項目 42-1 之後，發音控制指令

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

裝

## 五、發明說明(17)

剖析器 22 會將聲音控制指令 110 中發音參數 115 所定義的參數存入項目 42-1 的欄位 42-12, 42-13 及 42-14。圖 4A 中的 PID 欄位 42-11 是表示參數識別碼，此欄位是為了說明方便，在實際使用時可以忽略不用。

如圖 4B 所示，發音控制指令剖析器 22 在分析音效控制指令 120 時，會在音效表 43 中，加入或修改項目 43-1。首先發音控制指令剖析器 22 必須在音效表 43 取得可以使用的項目 43-1。項目 43-1 的取得方式可以是在音效表 43 中插入新的項目，或是重覆使用音效表 43 中不再使用的項目。在取得項目 43-1 之後，發音控制指令剖析器 22 會將音效控制指令 120 中的音效資料檔名 125 及其音效資料內容存入項目 43-1 的欄位 43-12 及 43-13。圖 4B 中的 AID 欄位 43-11 是表示音效資料識別碼，此欄位是為了說明方便，在實際使用時可以忽略不用。另外，為了節省記憶體空間之使用，發音控制指令剖析器 22 在修改音效表 43 之前，可使用音效資料檔名 125 對音效表 43 進行檢索，若發現相同的檔名已存在，則不做修改的動作。

發音控制指令剖析器 22 在修改參數表 42 或音效表 43 完成之後，接著修改標籤對應表 41 的內容。首先發音控制指令剖析器 22 以發音控制指令 110 或 120 的識別碼 111 或 121、超媒體標籤 113 或 123 及屬性欄位檢索標籤對應表 41，並獲得項目 41-1 或 41-2。在檢索中，若項目 41-1 或 41-2 不存在，則建立新的項目 41-1 或 41-2。然後發音控制指令剖析器 22 會針對聲音控制指令 110，

(請先閱讀背面之注意事項再填寫本頁)

裝

訂



## 五、發明說明(14)

將項目 41-1 的型態欄位 41-13 設定為 PARAM，以標明指標欄位 41-14 是指向參數表 42。同時發音控制指令剖析器 22 也將指標欄位 41-14 設定為該聲音控制指令 110 在參數表 42 中的對應項目 42-1。針對音效控制指令 120，發音控制指令剖析器 22 將項目 41-2 的型態欄位 41-23 設定為 AUDIO，以標明指標欄位 41-24 是指向音效表 43。同時發音控制指令剖析器 22 也將指標欄位 41-24 設定為該音效控制指令 120 在音效表 43 中的對應項目 43-1。

在圖 3 中，指令 130、140 及 150 是破音字控制指令，這些指令的識別碼 131、141 及 151 均為 "ALT"。每個破音字控制指令 130、140 及 150 均定義破音字串 133、143 及 153，以及取代破音字串 133、143 及 153 的替代字串 135、145 及 155。其目的是使用替代字串 135、145 及 155 讓文字語音轉換器 27 產生正確的發音。指令 140 及 150 也定義前後文字串 147 及 157，用來限制指令 140 及 150 所適用的範圍。

如圖 5A 所示，發音控制指令剖析器 22 在分析破音字控制指令 130、140 及 150 時，會在破音字表 31 中，依序加入或修改項目 31-1、31-2 及 31-3。發音控制指令剖析器 22 將破音字串 133、143 及 153 讀入並作適當的文字轉換，再分別存入項目 31-1、31-2 及 31-3 的破音字串欄位 31-11、31-21 及 31-31。發音控制指令剖析器 22 將替代字串 135、145 及 155 讀入並作適當的文字轉換，再分別存入項目 31-1、31-2 及 31-3 的替代字串欄位 31-12、31-22 及 31-32。發音控制指令剖析器 22 將前後文

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

## 五、發明說明(15)

字串 147 及 157 讀入並作適當的文字轉換，再分別存入項目 31-2 及 31-3 的前後文字串欄位 31-23 及 31-33。

在圖 3 中，指令 160，170 及 180 是專有名詞控制指令，這些指令的識別碼 161，171 及 181 均為“TERM”。每個專有名詞控制指令 160，170 及 180 均定義專有名詞字串 163，173 及 183，以及用來取代專有名詞字串 163，173 及 183 的譯文字串 165，175 及 185。其目的是使用譯文字串 165，175 及 185，讓文字語音轉換器 27 能夠針對專有名詞字串 163，173 及 183，產生正確的發音。例如當文字語音轉換器 27 只能轉換中文文字時，專有名詞控制指令 160，170 及 180 可用於將英文或中英混雜的專有名詞轉換為中文語音訊號。

如圖 5B 所示，發音控制指令剖析器 22 在分析專有名詞控制指令 160，170 及 180 時，會在專有名詞表 32 中，依序加入或修改項目 32-1，32-2 及 32-3。發音控制指令剖析器 22 將專有名詞字串 163，173 及 183 讀入並作適當的文字轉換(詳述於後)，再分別存入項目 32-1，32-2 及 32-3 的專有名詞欄位 32-11，32-21 及 32-31。發音控制指令剖析器 22 將譯文字串 165，175 及 185 讀入並作適當的文字轉換(詳述於後)，再分別存入項目 32-1，32-2 及 32-3 的譯文欄位 32-12，32-22 及 32-32。

在圖 2 中，文字轉換器 26 接收並處理由超媒體標示語言剖析器 24 所分離產生的文字內容。文字轉換器 26 搜尋該文字內容，以決定是否有破音字表 31 內的破音字串及專有名詞表 32 內的專有名詞。若發現有破音

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

家

## 五、發明說明(16)

字串或專有名詞，文字轉換器 26 將依據音字表 31 或專有名詞表 32 的項目內容，執行字串替代。文字轉換器 26 將處理完成的文字內容，送交給文字語音轉換器 27。

標籤轉換器 25 接收並處理由超媒體標示語言剖析器 24 所分離產生的超媒體標籤。標籤轉換器 25 用此超媒體標籤檢索標籤對應表 41，以決定該超媒體標籤是否需要做聲音控制或是音效控制。若發現該超媒體標籤需要做聲音控制，標籤轉換器 25 依據檢索獲得之項目的指標欄位，自參數表 42 取得對應之參數，並將這些發音參數送至文字語音轉換器 27。若發現該超媒體標籤需要做音效控制，標籤轉換器 25 依據檢索獲得之項目的指標欄位，自音效表 43 取得對應之音效資料，並將其送至音效裝置 16 或是電話介面裝置 14。

文字語音轉換器 27 接收並處理來自文字轉換器 26 的文字內容，以及來自標籤轉換器 25 的發音參數。文字語音轉換器 27 會依據新收到的發音參數，改變其參數設定，例如改變聲音的速度、音量及韻律等參數之設定值。文字語音轉換器 27 接收到文字內容時，會依據當時發音參數之設定值，將該文字內容轉換為語音訊號，並將結果送至音效裝置 16 或是電話介面裝置 14。

圖 6 說明文件讀取控制器 28 的執行步驟。在步驟 S1，超媒體語音轉換系統 10（中央處理器 11 執行作業系統或應用程式）決定是否需要讀取獨立的發音控制指令檔 21。若是的話，文件讀取控制器 28 會執行步驟

(請先閱讀背面之注意事項再填寫本頁)

訂

線

## 五、發明說明(17)

S2，讀取該檔案的內容，並在步驟 S6 將檔案內容交給發音控制指令剖析器 22 分析發音控制指令。在執行完步驟 S6，或是不需要讀取獨立的發音控制指令檔 21，文件讀取控制器 28 執行步驟 S3，讀取超媒體文件檔案 23 的內容。在步驟 S4 中，文件讀取控制器 28 將超媒體文件檔案 23 的內容交由超媒體標示語言剖析器 24 分離成文件元件。一文件元件可以是一超媒體標籤、一文字內容的文字串或是一發音控制指令。文件讀取控制器 28 則依序讀取並處理超媒體標示語言剖析器 24 分離出來的文件元件。在步驟 S5 中，文件讀取控制器 28 決定超媒體標示語言剖析器 24 分離出來的文件元件是否為發音控制指令。若是的話，文件讀取控制器 28 執行步驟 S6，將該發音控制指令交由發音控制指令剖析器 22 分析。執行完步驟 S6 後，文件讀取控制器 28 回到步驟 S4，讀取並處理下一由超媒體標示語言剖析器 24 分離出來的文件元件。在步驟 S5 中，如果讀取的文件元件不是發音控制指令，文件讀取控制器 28 執行步驟 S7。如果文件讀取控制器 28 在步驟 S7 發現該文件元件是一超媒體標籤，文件讀取控制器 28 在步驟 S8 中，將該超媒體標籤交由標籤轉換器 25 進行聲音及音效的控制。然後文件讀取控制器 28 回到步驟 S4，讀取並處理下一文件元件。在步驟 S7 中，如果讀取的文件元件不是超媒體標籤，文件讀取控制器 28 執行步驟 S9。在步驟 S9 中，文件讀取控制器 28 將讀取的文件元件視為一文字內容的文字串，並將其交由文字轉換器

(請先閱讀背面之注意事項再填寫本頁)

訂

線

## 五、發明說明(18)

26 做文字的替代轉換。在步驟 S10 中，文件讀取控制器 28 把轉換之結果交給文字語音轉換器 27 處理，文字語音轉換器 27 會將其轉換為語音訊號，並經由音效裝置 16 或電話介面裝置 14 播放出來。然後文件讀取控制器 28 會回到步驟 S4，讀取並處理下一文件元件。文件讀取控制器 28 會重覆執行這些步驟 S4-S10，直到超媒體文件檔案 23 內所有的文件元件都處理完。

圖 7 說明發音控制指令剖析器 22 的執行流程。在步驟 S11 中，發音控制指令剖析器 22 讀取發音控制指令。在步驟 S12，發音控制指令剖析器 22 依據指令的識別碼，判別該發音控制指令是否為聲音控制指令。若是的話，發音控制指令剖析器 22 執行步驟 S13，在標籤對應表 41 中加入或修改一項目，並將該指令的超媒體標籤、指令型態(在此為 PARAM)、屬性及參數指標存入對應的欄位。然後發音控制指令剖析器 22 執行步驟 S14，將該指令所定義的發音參數存入參數表 42。執行完步驟 S14 後，發音控制指令剖析器 22 回到步驟 S11，讀取並處理下一發音控制指令。

在步驟 S12，如果指令不是聲音控制指令，發音控制指令剖析器 22 執行步驟 S15，判別該指令是否為音效控制指令。若是的話，發音控制指令剖析器 22 執行步驟 S16，在標籤對應表 41 中加入或修改一項目，並將該指令的超媒體標籤、指令型態(在此為 AUDIO)、屬性、音效資料指標存入對應的欄位。然後發音控制指令剖析器 22 執行步驟 S25，將該指令所定義的音效檔名

(請先閱讀背面之注意事項再填寫本頁)

訂

線

## 五、發明說明(19)

及其檔案內容存入音效表 43。執行完步驟 S25 後，發音控制指令剖析器 22 回到步驟 S11，讀取並處理下一發音控制指令。

在步驟 S15，如果指令不是音效控制指令，發音控制指令剖析器 22 執行步驟 S17，判別該指令是否為專有名詞控制指令。若是的話，發音控制指令剖析器 22 執行步驟 S18，依據破音字表 31 現有的內容，對該指令進行文字轉換。也就是說，針對破音字表 31 的各個項目，依序檢查該指令是否存在有需要轉換的破音字串。若在該指令中發現需要轉換的破音字串，發音控制指令剖析器 22 則將其轉換為對應的替代字串。執行完步驟 S18，發音控制指令剖析器 22 執行步驟 S19，依據專有名詞表 32 現有的內容，對該指令進行文字轉換。也就是說，針對專有名詞表 32 的各個項目，依序檢查該指令是否存在有需要轉換的專有名詞字串。若在該指令中發現需要轉換的專有名詞字串，發音控制指令剖析器 22 則將其轉換為對應的譯文字串。執行完步驟 S19，發音控制指令剖析器 22 執行步驟 S20，將轉換後的專有名詞控制指令存入專有名詞表 32。然後發音控制指令剖析器 22 回到步驟 S11，讀取並處理下一發音控制指令。

在步驟 S17，如果指令不是專有名詞控制指令，發音控制指令剖析器 22 執行步驟 S21，判別該指令是否為破音字控制指令。若是的話，發音控制指令剖析器 22 執行步驟 S22，依據破音字表 31 現有的內容，對該指

(請先閱讀背面之注意事項再填寫本頁)

訂

線

## 五、發明說明(20)

令進行文字轉換。也就是說，針對破音字表 31 的各個項目，依序檢查該指令是否存在有需要轉換的破音字串。若在該指令中發現需要轉換的破音字串，發音控制指令剖析器 22 則將其轉換為對應的替代字串。執行完步驟 S22，發音控制指令剖析器 22 執行步驟 S23，將轉換後的破音字控制指令存入破音字表 31。然後發音控制指令剖析器 22 回到步驟 S11，讀取並處理下一發音控制指令。

在步驟 S21，如果指令不是破音字控制指令，發音控制指令剖析器 22 執行步驟 S24。在步驟 S24 中，發音控制指令剖析器 22 將所讀取的資料視為註解，因此將其忽略。然後發音控制指令剖析器 22 回到步驟 S11，讀取並處理下一發音控制指令。發音控制指令剖析器 22 會重覆執行步驟 S11-S24，直到所有的發音控制指令都處理完。

圖 8 說明文字轉換器 26 的執行流程。在步驟 S31 中，文字轉換器 26 讀取超媒體文件檔案 23 的文字內容。接著文字轉換器 26 執行步驟 S32，依據破音字表 31 現有的內容，對該文字內容進行文字轉換。也就是說，文字轉換器 26 會針對破音字表 31 的各個項目，依序檢查該文字內容是否存在有需要轉換的破音字串。若在該文字內容中發現需要轉換的破音字串，文字轉換器 26 則將其轉換為對應的替代字串。為了執行上的效率，文字轉換器 26 在針對破音字表 31 的一項目做文字轉換時，首先搜尋該文字內容，找出該項目的破音字串的位

(請先閱讀背面之注意事項再填寫本頁)

表

訂

像

## 五、發明說明(2)

置。再以該位置的前後字元或字串，加上該破音字串，與該項目的前後文字串進行比對，以決定該文字內容的這個破音字串是否需要以該項目的替代字串來取代。若是的話，文字轉換器 26 就以該項目的替代字串來取代這個破音字串。如果該項目沒有定義前後文字串，文字轉換器 26 則直接以替代字串來取代這個破音字串。處理完後，文字轉換器 26 繼續以此方式搜尋該文字內容，找出並處理下一該項目的破音字串，直到該文字內容都搜尋完畢。

執行完步驟 S32，文字轉換器 26 接著執行步驟 S33，依據專有名詞表 32 現有的內容，對步驟 S32 所轉換完成的文字內容進行文字轉換。也就是說，文字轉換器 26 會針對專有名詞表 32 的各個項目，依序檢查該指令是否存在有需要轉換的專有名詞字串。若在該文字內容中發現需要轉換的專有名詞字串，文字轉換器 26 則將其轉換為對應的譯文字串。文字轉換器 26 在針對專有名詞表 32 的一項目做文字轉換時，首先搜尋該文字內容，找出該項目的專有名詞字串，並直接以譯文字串來取代這個專有名詞字串。處理完後，文字轉換器 26 繼續以此方式搜尋該文字內容，找出並處理下一該項目的專有名詞字串，直到該文字內容都搜尋完畢。

執行完步驟 S33，文字轉換器 26 回到步驟 S31，讀取並處理下一串由超媒體標示語言剖析器 24 所產生的文字內容。

(請先閱讀背面之注意事項再填寫本頁)

衣

訂

線



## 五、發明說明(22)

圖 9 說明標籤轉換器 25 的執行步驟。為了處理超媒體標籤的巢狀式排列方式，標籤轉換器 25 使用一儲存在主記憶體 12 或中央處理器 11 中的堆疊(stack)，以便於執行超媒體標籤的轉換處理。在步驟 S41，針對一由超媒體標示語言剖析器 24 所產生的超媒體標籤，標籤轉換器 25 首先判別其是否為起始標籤。若該超媒體標籤是一起始標籤，標籤轉換器 25 會執行步驟 S42，將其推進(push)堆疊。否則的話，標籤轉換器 25 會執行步驟 S43，自堆疊中彈出(pop)一超媒體標籤。

在步驟 S44，標籤轉換器 25 針對堆疊頂端的超媒體標籤進行標籤轉換，並以該超媒體標籤檢索標籤對應表 41。在步驟 S45，標籤轉換器 25 依據檢索的結果，決定該超媒體標籤是否有對應的參數設定項目(其型態欄位為 PARAM)。若有的話，標籤轉換器 25 執行步驟 S46，使用該項目的參數指標，自參數表 42 中讀取對應的參數。然後標籤轉換器 25 將這些參數送交給文字語音轉換器 27，以改變往後文字內容的發音方式。

執行完步驟 S46，或是該超媒體標籤沒有對應的參數設定項目，標籤轉換器 25 執行步驟 S47，依據檢索的結果，決定該超媒體標籤是否有對應的音效控制項目(其型態欄位為 AUDIO)。若有的話，標籤轉換器 25 執行步驟 S48，使用該項目的音效資料指標，自音效表 43 中讀取對應的音效資料。然後標籤轉換器 25 將此音效資料送交給音效裝置 16 或電話介面裝置 14 播出。

(請先閱讀背面之注意事項再填寫本頁)

訂

線

經濟部中央標準局員工消費合作社印製

## 五、發明說明(23)

在步驟 S47，如果該超媒體標籤沒有對應的音效控制項目，標籤轉換器 25 執行步驟 S49，將該超媒體標籤忽略不處理。執行完步驟 S48 或步驟 S49，標籤轉換器 25 回到步驟 S41，等待處理下一由超媒體標示語言剖析器 24 所產生的超媒體標籤。

標籤轉換器 25 所使用的堆疊，可以確保內層的超媒體元件(HTML element)可使用自己的聲音及音效控制。同時當回到上層的超媒體元件時，仍能恢復該元件所使用的聲音及音效控制。

上述的實施例，只是為了說明本發明所提出的方法。熟悉本行技藝之人士，尚可導出不同的實施方式，而離不開下列申請專利範圍所揭露的精神與範圍。

(請先閱讀背面之注意事項再填寫本頁)

訂  
線

## 六、申請專利範圍

### 1. 一種將超媒體文件轉換為語音訊號的電腦系統，包括：

一超媒體標示語言剖析器，將一超媒體標示語言格式的文件，分離成文字內容、標示文件結構的超媒體標籤及控制發音方式的發音控制指令；

一發音控制指令剖析器，分析該發音控制指令，並依據該發音控制指令的內容修改標籤對應表、音效表、參數表、破音字表及專有名詞表；

一文字轉換器，針對該破音字表內規定必須修改發音的字串，在該文字內容中出現時，就依照該破音字表內所規定的方式，修改發音，而且針對該專有名詞表內規定必須翻譯的字串，在該文字內容中出現時，就依照該專有名詞表內所規定的方式做翻譯；

一標籤轉換器，針對該標籤對應表內規定必須修改發音參數或插入音效的超媒體標籤，在該超媒體文件中出現時，就依照該標籤對應表內所指定的該音效表的項目，插入音效資料，並且使用該標籤對應表內所指定的該參數表的項目，修改該超媒體標籤所標示的該文字內容的發音參數；以及

一文字語音轉換器，將該文字內容經由該文字轉換器及該標籤轉換器修改後的結果，轉換為語音訊號。

### 2. 一種在一發音控制指令剖析器中將超媒體文件轉換為語音訊號的方法，包括下列步驟：

剖析指定特定超媒體標籤應使用的音量、速度及韻律等發音參數的聲音控制指令，及剖析指定特定超

(請先閱讀背面之注意事項再填寫本頁)

訂

## 六、申請專利範圍

### 1. 一種將超媒體文件轉換為語音訊號的電腦系統，包括：

一超媒體標示語言剖析器，將一超媒體標示語言格式的文件，分離成文字內容、標示文件結構的超媒體標籤及控制發音方式的發音控制指令；

一發音控制指令剖析器，分析該發音控制指令，並依據該發音控制指令的內容修改標籤對應表、音效表、參數表、破音字表及專有名詞表；

一文字轉換器，針對該破音字表內規定必須修改發音的字串，在該文字內容中出現時，就依照該破音字表內所規定的方式，修改發音，而且針對該專有名詞表內規定必須翻譯的字串，在該文字內容中出現時，就依照該專有名詞表內所規定的方式做翻譯；

一標籤轉換器，針對該標籤對應表內規定必須修改發音參數或插入音效的超媒體標籤，在該超媒體文件中出現時，就依照該標籤對應表內所指定的該音效表的項目，插入音效資料，並且使用該標籤對應表內所指定的該參數表的項目，修改該超媒體標籤所標示的該文字內容的發音參數；以及

一文字語音轉換器，將該文字內容經由該文字轉換器及該標籤轉換器修改後的結果，轉換為語音訊號。

### 2. 一種在一發音控制指令剖析器中將超媒體文件轉換為語音訊號的方法，包括下列步驟：

剖析指定特定超媒體標籤應使用的音量、速度及韻律等發音參數的聲音控制指令，及剖析指定特定超

(請先閱讀背面之注意事項再填寫本頁)

訂

## 六、申請專利範圍

媒體標籤應使用的音效資料的音效控制指令，而且根據該聲音控制指令的內容，將發音參數存入參數表項目中，並在標籤對應表中，設定該超媒體標籤與該參數表項目的對應關係，以及根據該音效控制指令的內容，將音效資料存入音效表項目中，並在該標籤對應表中，設定該超媒體標籤與該音效表項目的對應關係；以及

剖析指定超媒體文件內文字內容中特定字串的發音修改方式及前後文字串的破音字控制指令，及剖析指定該文字內容中特定字串的翻譯方式的專有名詞控制指令，其中該發音修改方式及該翻譯方式是可由一文字語音轉換器轉換為語音訊號的文字串，而且該發音控制指令剖析器根據該破音字控制指令的內容，產生破音字表，存放該特定字串及其對應的該發音修改方式及該前後文字串，以及根據該專有名詞控制指令的內容，產生專有名詞表，存放該特定字串及其對應的該翻譯方式。

3. 一種在一文字轉換器中將超媒體文件轉換為語音訊號的方法包括下列步驟：

針對破音字表各個項目所規定必須修改發音的特定字串，將超媒體文件的文字內容中的該特定字串，以其指定的用於修改發音的替代字串取代，該替代字串可以將該特定字串的多種發音，指定為其中一種特定的發音；以及

針對專有名詞表各個項目所規定必須作翻譯的

(請先閱讀背面之注意事項再填寫本頁)

訂

## 六、申請專利範圍

特定字串，將該超媒體文件的文字內容中的該特定字串，以其指定的用於翻譯的譯文字串取代，該譯文字串可以讓該特定字串中部份無法轉換為語音訊號的字串，經由文字語音轉換器轉換為指定的語音訊號。

4. 一種在一標籤轉換器中將超媒體文件轉換為語音訊號用以修改發音參數及插入音效的方法包括下列步驟：

針對標籤對應表內各個項目所規定必須修改發音參數的超媒體標籤，將超媒體文件中該超媒體標籤所標示的文字內容，依照該標籤對應表項目所指定的參數表項目，修改音量、速度及韻律等發音參數；以及

針對該標籤對應表內各個項目所規定必須插入音效的超媒體標籤，對該超媒體文件中的該超媒體標籤，依照該標籤對應表項目所指定的音效表項目，產生音效。

5. 一種將超媒體文件轉換為語音訊號的方法包括下列步驟：

剖析發音控制指令，這個步驟包括下列步驟：

針對聲音控制指令，產生一由該聲音控制指令指定的超媒體標籤來檢索的標籤對應表項目，以及產生一參數表項目，儲存該聲音控制指令所指定的音量、速度及韻律等發音參數，並且在該標籤對應表項目中，設定其指標欄位內的指標指向該參數表項目；

(請先閱讀背面之注意事項再填寫本頁)

訂

## 六、申請專利範圍

針對音效控制指令，產生一由該音效控制指令指定的超媒體標籤來檢索的標籤對應表項目，以及產生一音效表項目，儲存該音效控制指令所指定的音效資料，並且在該標籤對應表項目中，設定其指標欄位內的指標指向該音效表項目；

針對破音字控制指令，產生一由該破音字控制指令指定必須修改發音的特定字串來檢索的破音字表項目，儲存該破音字控制指令所指定的替代字串，這個替代字串可以將具有多種發音的該特定字串，轉換為其中一種特定的發音；以及

針對專有名詞控制指令，產生一由該專有名詞控制指令指定必須作翻譯的特定字串來檢索的專有名詞表項目，儲存該專有名詞控制指令所指定的譯文字串，這個譯文字串可以讓原本無法由文字語音轉換器轉換為語音訊號的該特定字串，轉換為特定的語音訊號。

6. 如申請專利範圍第 5 項所述之方法，進一步包括下列步驟：從該超媒體文件的一註解中萃取該發音控制指令。

7. 如申請專利範圍第 5 項所述之方法，進一步包括下列步驟：讀取各個該發音控制指令。

(請先閱讀背面之注意事項再填寫本頁)

訂

## 六、申請專利範圍

8. 如申請專利範圍第 5 項所述之方法，進一步包括下列步驟：

分析一起媒體文件的資料；

遇到一起媒體標籤時，就以該超媒體標籤檢索該標籤對應表；

針對檢索而得的該標籤對應表項目，使用其指標欄位內的指標，取得該參數表項目及該音效表項目；

使用該參數表項目內的一組發音參數，修改往後文字語音轉換器所使用的發音參數；以及

插入該音效表項目所存放的音效資料。

9. 如申請專利範圍第 8 項所述之方法，進一步包括下列步驟：

遇到一起媒體標籤中的起始標籤時，將該起始標籤推入堆疊；以及

遇到一起媒體標籤中的結束標籤時，就從該堆疊中彈出一起媒體標籤，

其中用於檢索之該特定超媒體標籤，就是位於該堆疊頂端的一起媒體標籤。

10. 如申請專利範圍第 5 項所述之方法，進一步包括下列步驟：

分析一起媒體文件的資料；

搜尋該超媒體文件的文字內容；

(請先閱讀背面之注意事項再填寫本頁)

裝

訂



## 六、申請專利範圍

將該文字內容中，凡是符合該破音字表內規定必須修改發音的字串，就以該破音字表項目內所指定的替代字串取代；以及

將該文字內容中，凡是符合該專有名詞表內規定必須作翻譯的字串，就以該專有名詞表項目內所指定的譯文字串取代。

11. 如申請專利範圍第 10 項所述之方法，其中該破音字表項目還含有一前後文欄位，該文字內容中符合該破音字表項目內規定必須修改發音的字串，其前後文也必須符合該項目的前後文欄位內所規定的前後文字串，在此情況下，該字串才會以該破音字表項目內所指定的替代字串取代。

12. 如申請專利範圍第 11 項所述之方法，進一步包括下列步驟：

產生一語音訊號，其內容包括該音效資料所產生的語音訊號，以及由該超媒體文件的該文字內容、該替代字串及該譯文字串，依照該發音參數所轉換完成的語音訊號。

(請先閱讀背面之注意事項再填寫本頁)

訂

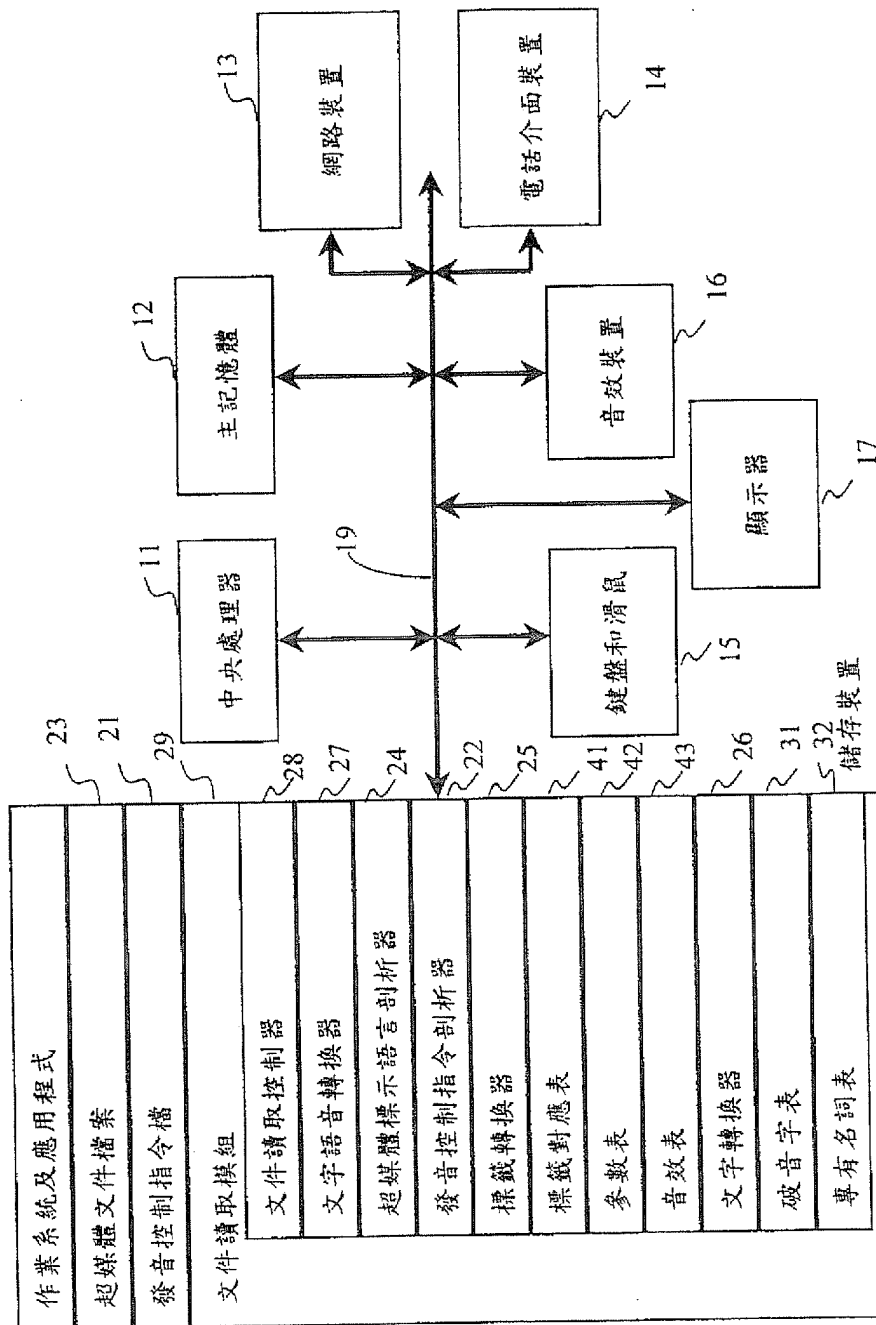


圖 1

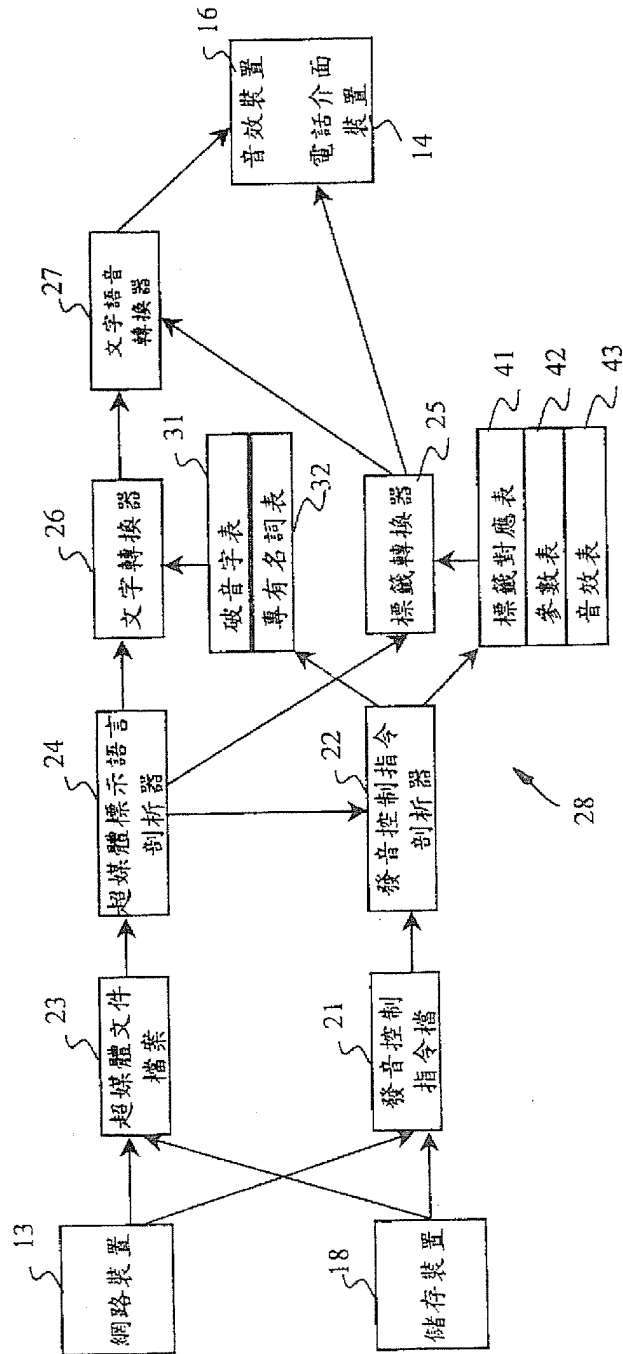


圖 2

# <!-- HTS Control rules

```

110  {PARAM LI <speed=1.0,volume=0.8,pitch=1.2>
      111  113  115
120  {AUDIO LI beep.au
      121  123  125
130  {ALT 率緣
      131  133  135
140  {ALT 率帥率領
      141  143  145  147
150  {ALT 爲維轉換爲
      151  153  155  157
160  {TERM web 全球資訊網
      161  163  165
170  {TERM "world wide web" 全球資訊網
      171  173  175
180  {TERM HTML文件 超媒體文件
      181  183  185
-->

```

圖 3

PID	速度	音量	韻律
0	1.0	0.8	1.2
...	...	...	...

42-1

42-11 42-12 42-13 42-14

42

圖 4A

AID	檔名	音效資料
0	beep.au	...
...	...	...

43-1

43-11 43-12 43-13

43

圖 4B

標籤	屬性	型態	指標
L1		PARAM	0
L1		AUDIO	0
...	...	...	...

41-1

41-2

41-11 41-12 41-13 41-14 41-21 41-22 41-23 41-24

41

圖 4C

31-11	31-12	31-13
破音字串	替代字串	前後文字串
率	綠	
綠	帥	綠領
爲	維	轉換爲
...	...	...

31-1 31-2 31-3

31 31-21 31-22 31-23 31-31 31-32 31-33

圖 5A

32-11	32-12
專有名詞	譯文
web	全球資訊網
world wide 全球資訊網	全球資訊網
HTML 文件	超媒體文件
...	...

32-1 32-2 32-3

32 32-21 32-22 32-31 32-32

圖 5B

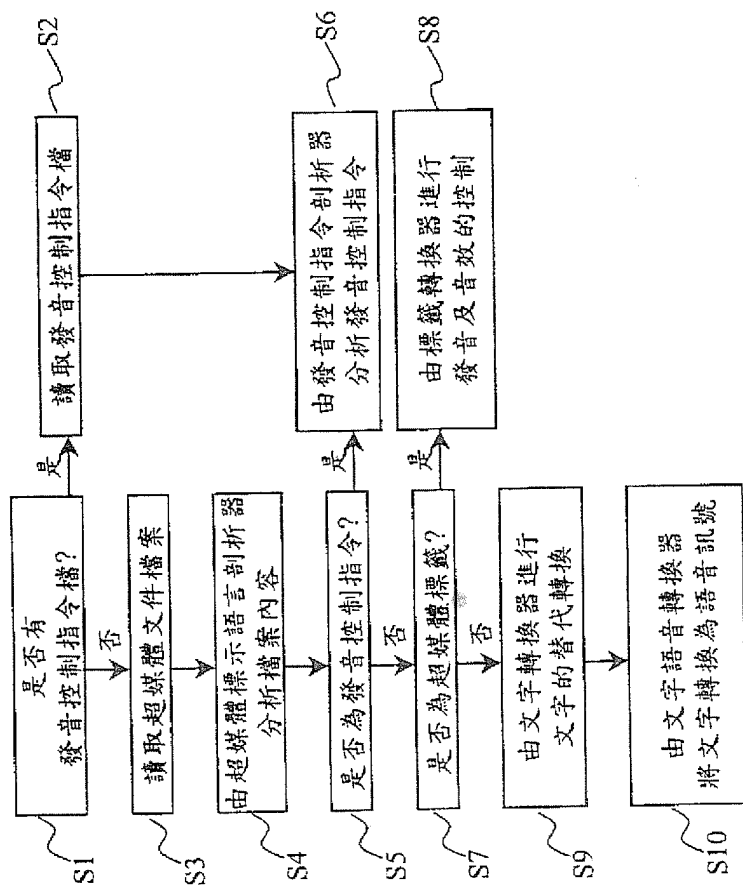


圖 6

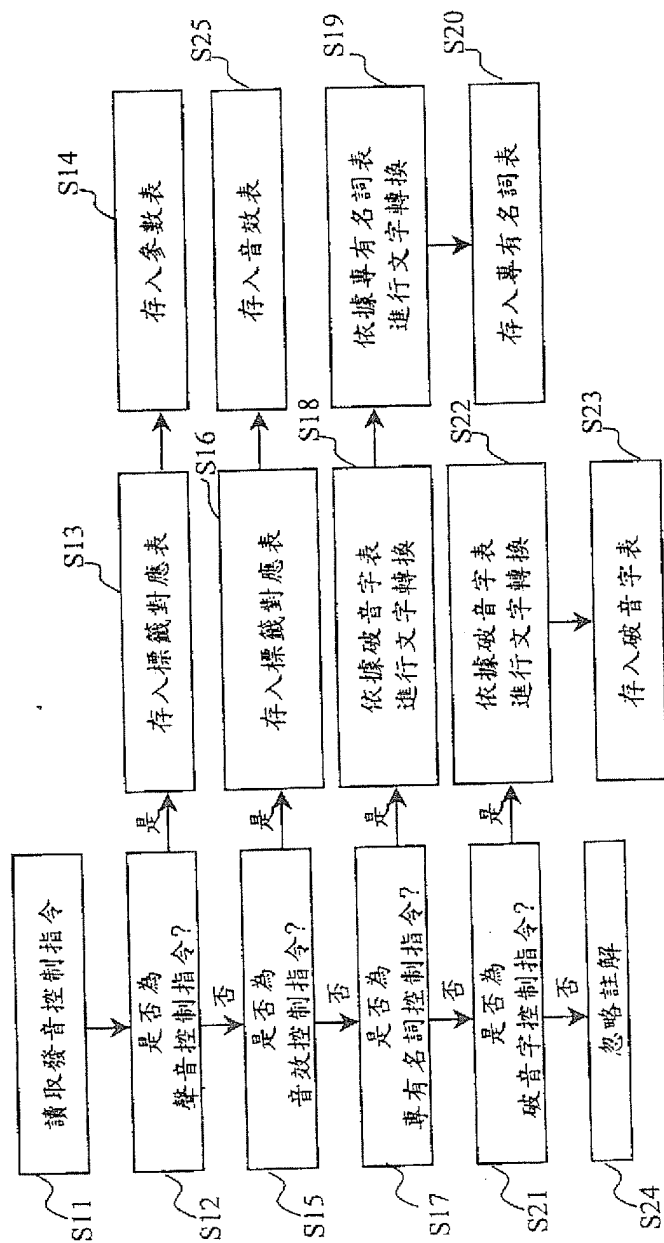


圖 7



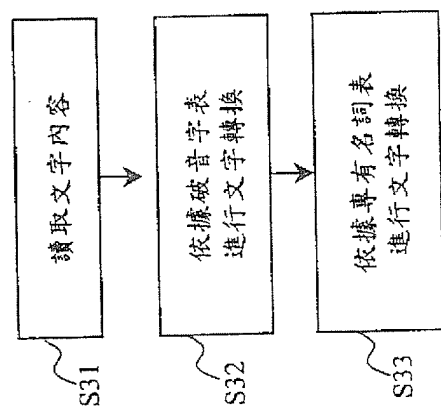


圖 8

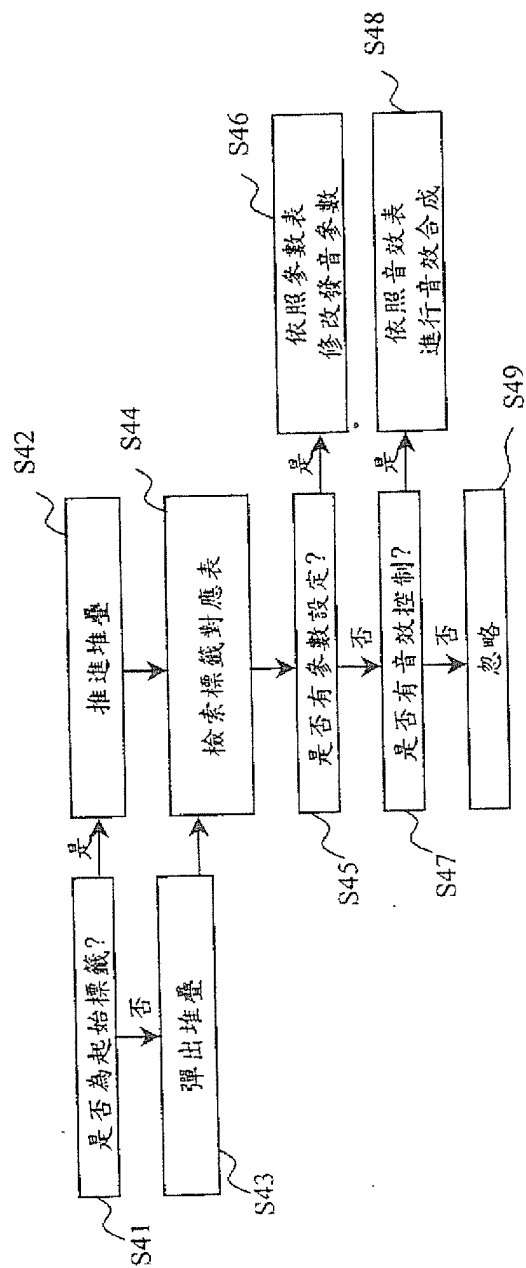


圖 9